

Chronology Compilation Tool (CCT) for Analyzing Service Patterns

Don Schroeder, Ph.D.

Carol Forhan, MBA

Jon D. Busch, Ph.D.

Thomson Medstat

**SAMHSA Integrated Medicaid and State Agency Mental Health and Substance
Abuse Data Analysis and Technical Assistance Project**

Contract No. 270-01-7087

March 2007

Chronology Compilation Tool (CCT) for Analyzing Service Patterns

Introduction

Chronology Compilation Tool (CCT) is an analytic approach that can help health service researchers, statisticians, and programmers work with complex administrative health claims and encounter databases. CCT can be used to analyze health care treatment patterns by clients over discreet periods of time to better understand the care provided and related outcomes. This paper explains the analytic tool and gives examples of how this tool was used in research studies using the SAMHSA Integrated Database (IDB).

Research studies that investigate mental health (MH) and/or substance abuse (SA) service utilization are often based on annualized rates of specified categories of service. These studies may use data from administrative claims and encounters systems to examine services across a study population. In this type of study, a useful statistic may be the proportion of a MH study population that received between five to ten days of inpatient MH treatment during a year. While such high-level analyses can provide important information, researchers may want to explore and understand variations in care for the subpopulation of clients who receive treatment. Specifically, examination of treatment patterns within a defined timeframe can provide insight as to the types of care being provided and the impact on client outcomes and future service utilization.

In the example noted above, a client might receive inpatient MH services ten times in a year. These visits could have taken place in a month, over a period of three months, or over a longer period. Understanding a more exact window of treatment allows researchers to identify patterns in service use. In addition, the researcher then has the ability to collapse or eliminate possible duplicative services reported by multiple agencies, addressing an issue that frequently arises when multiple data sources are used. These temporal analyses involve “drilling down” to examine trends and patterns in services using a finer level of granularity in terms of timeline, such as analysis of services by quarter, month, or week within a year. The window of time is defined by the researcher and can be used iteratively to explore the data.

Researchers using the SAMHSA Integrated Database (IDB) became interested in building and using a more flexible research tool as they began to look more closely at service utilization patterns for clients receiving MH and/or SA services from one or more state public agencies. The IDB includes claims and encounter data from three states – Delaware, Oklahoma, and Washington. These data are obtained from each state’s Medicaid agency and also from state MH and state SA agencies. In order to analyze this information, the researchers needed to find a way to summarize and organize data around meaningful windows of time – in other words, creating temporal analyses. These analyses have been used in support of research to investigate utilization of services across all agencies, to examine co-occurring mental health and substance abuse conditions, and to understand the impact of healthcare treatment received by clients following detoxification services. This report describes the Chronology Compilation Tool (CCT) and provides examples of its application and utility.

Fixed and Sliding Timeframes

A basic example of temporal analyses is the examination of the pattern of inpatient hospitalizations by quarter for a given year. Treated patients can be hospitalized in one or more quarters and a finite number of treatment patterns can be defined. The table below illustrates the 15 possible combinations for a given individual. Each “I” indicates a quarter in which the client received inpatient services. It is interesting to note that at a high level, each of these patterns describes a person who received services within a year, although a client classified in “pattern 1” almost certainly uses different levels and types of resources than a patient classified in “pattern 15.”

A fixed timeframe is a defined period of time used for analysis – for example a calendar quarter – January 1 through March 31.

A sliding timeframe is a period of time that may be unique for each data unit of analysis. In an analysis of healthcare claims, the sliding timeframe would begin and end based on an individual’s services rather than a fixed calendar period.

TABLE 1: POSSIBLE COMBINATIONS OF QUARTERLY TREATMENT PROFILES FOR A SINGLE CLIENT

Pattern	Q1	Q2	Q3	Q4
1	I			
2		I		
3			I	
4				I
5	I	I		
6	I		I	
7	I			I
8		I	I	
9		I		I
10			I	I
11	I	I	I	
12	I	I		I
13	I		I	I
14		I	I	I
15	I	I	I	I

We can represent these patterns using a reporting shorthand as follows: I---, -I--, --I-, ---I, II--, I-I-, I--I, -II-, -I-I, --II, III-, II-I, I-II, -III, IIII, where “I” represents a hospitalization within a quarter and “-” indicates no hospitalization for that quarter.

Since we can classify all patients into one of 15 groups, this basic analysis can be used to examine variation in utilization and outcomes among the various groups. This example is considered a **“fixed” timeframe** because the demarcation points are somewhat arbitrary quarterly boundaries. Using the fixed timeframes above, a patient who was hospitalized once on 9/20 and again on 10/5 would be classified into group 10, whereas a patient with inpatient services on 9/20 and 9/25 would be classified into group 3.

In order to minimize the impact of classification based on arbitrary boundaries, we can construct 90-day “windows” around a given event such as an initial hospitalization; this would constitute a **“sliding” timeframe**.

Returning to the original example of the patient who received inpatient services on 9/20, we would construct four 90-day windows based on an initial start date of 9/20. We would then look for services in each of these windows and construct a treatment profile based on classification into one of 15 possible groups.

Since these timeframes are not based on fixed dates (such as 1/1, 4/1, 7/1, and 10/1), these windows “slide” from person to person based on the initial event. This requires us to construct literally hundreds of analytic windows – a task that is difficult when performed manually, but can be accomplished rapidly using current software and hardware tools.

Sequential Analyses

Each of the examples described thus far illustrates how time can be used as a dimension to study a single type of service. We can adapt this general approach to examine the sequence of multiple types of settings for services (e.g., inpatient, residential, and outpatient), types of services (e.g., MH, SA) or provider/payer (e.g., Medicaid, State Agency) and to determine whether these situations were associated with enrollment in Medicaid or another specific program.

As an example, we could investigate whether patients had received an inpatient service and/or an outpatient service within a given 90-day period (fixed or sliding). Further, we could examine whether patients with one type of service subsequently received another type of service.

One special study using the IDB data examined quarterly profiles of both inpatient (I) and outpatient (O) SA treatments for clients. These profiles were used to determine the proportion of the population that had outpatient SA treatment in the quarter following a quarter in which inpatient SA treatment was received. In this study, understanding the patterns of inpatient and outpatient care helped researchers analyze the rate of appropriate follow-up care.

Sequential analyses of service patterns across small intervals of time (e.g., day, week, or month) can provide the means of defining treatment episodes and recurrence rates of medical problems.

For example, an IDB-based study of repeated detoxification episodes focused on analyzing daily service patterns over a sliding (not static) time frame.

Development of Person-Level Measures

The Chronology Compilation Tool utilizes a programming language, such as SAS, to construct a client-level summary of daily service characteristics over an entire year, or even multiple years, into a single [string] variable called a client-timeline (CTL). In addition to defining daily service characteristics, the CTL can also capture a daily program enrollment status for each client.

The CTL variable by itself has limited value because of its length. However, the variable can be used as a base for compiling utilization and enrollment statistics into static time periods (e.g., month, quarter, and year) or into service episodes.

The tool has generally been used to summarize and compile parameters of type of service (e.g., MH and/or SA and/or other medical service), setting of service (e.g., inpatient and/or residential and/or outpatient/other), reporting agency (e.g., Medicaid and/or other state agency), and Medicaid enrollment status (e.g., enrolled - not disabled, enrolled - disabled, or not enrolled). This tool can be modified to summarize other service or enrollment parameters appropriate to a particular study.

Efficiencies and Limitations

It is worth noting that this approach can be relatively complex in terms of data preparation and can be difficult to implement on a one-time basis. This is offset by the ability to easily conduct analyses using a wide range of time parameters. For example, after raw data have been loaded and the basic CCT routines applied, it is possible to re-analyze the same data using quarterly, monthly, or even weekly timelines.

The CCT routines were originally developed to prevent over-counting of services that may be reported from multiple data sources. This is typically accomplished by examining all service dates for each individual and coding each day either as a day on which services were provided or a day on which no specific services were provided. In this example services are categorized by type of service, setting of service, and agency providing the services. While this works with semi-additive events such as services, the same approach does not apply to additive values, such as costs. That is, it is not always possible to determine whether two costs represent “duplicate” costs for the same service, separate costs for different services, or separate shares of costs for the same service.

Examples of How the Tool Was Used

The Chronology Compilation Tool has often been used to develop client-level qualitative and quantitative measures, based on daily service profiles, for Medicaid enrollment status (E), type of service (T), setting of service (S) and reporting agency (R). Each [daily] segment of the entire CTL string collected some permutation of values for each of these parameters. This taxonomy is outlined in Table 2, and a partial enumeration of the 72 possible values ($3 * 4 * 3 * 2$) for a three-year, quarterly study is included in Table 3. For a three-year, quarterly study, these would be further indexed by indicators for each of the 12 quarters, resulting in 864 possible combinations.

For one study, researchers examined Medicaid eligibility by applying the CCT algorithm to summarized service usage over multiple, overlapping time periods for clients with MH and/or SA services (Federman et al., 2007). The time periods were of varying lengths, with differing begin dates and end dates. The high-level period identified a MH/SA “episode” or service window. A mid-level period spanned from the earliest to the last enrollment date in the service window. The most detailed periods of the study marked periods of Medicaid enrollment (and non-enrollment); enrollment periods separated by less than 30-days were combined. This mid-level enrollment span might consist of a single, uninterrupted Medicaid enrollment period or several enrollment periods (separated by enrollment gaps) may comprise the span. Using these overlapping time periods, we easily summarized usage across several categories such as number of Medicaid enrollment periods, setting of services, type of service, and client demographics.

Service utilization patterns for individuals with serious mental illnesses by client and provider characteristics were examined in another study (Buck 2001). In the initial phase, episodes of care were defined and measures of utilization were compiled for each episode. Researchers used an algorithm [one module of the CCT] that navigated through each client timeline to define episodes of care. This ‘step-through’ algorithm can be termed a ‘crawler’ code, one that interprets patterns along each client timeline. The lengthy process of processing all of the service-level records to construct summary client timelines was conducted only once. It was possible to save and reuse the client timelines in successive runs of refined crawler code. Similarly, the episode evaluations could be performed with either a sliding window perspective or a static window perspective with few changes to the crawler code. Overall, the CCT afforded great flexibility to the analysts, as they were able to pursue various analytic approaches with minimal costs related to computing time and programming effort.

For a study researching the factors affecting detoxification readmission (Mark et al., 2006), researchers developed timelines for all clients having at least one detoxification event within a three-year period. Initially, episodes of post-detoxification services were defined and episode measures were compiled. The nature of the algorithm made it possible to quickly recompile statistics based on revised episode definitions that utilized previously developed client timelines in less time than the long processing times necessary to develop statistics from service-level records,

Table 2: Typical parameters used in CTL analyses

Enrollment in Medicaid		Type of Service		Setting of Service		Reporting Agency
Enrolled (not disabled)	by	Mental Health	by	Inpatient	by	Agency (state)
Disabled (and enrolled)		Substance Abuse		Residential		
Not enrolled		Integrated Program		Outpatient	Medicaid	
		Other (medical)				

**TABLE 3: PARTIAL LIST OF POSSIBLE PARAMETERS FOR A CTL STUDY
SPANNING THREE YEARS (12 QUARTERS)**

E (Enrollment)	T (Treatment)	S (Setting)	R (Reporting)	Q (Quarter)
E	M	I	A	1
E	M	I	M	1
E	M	R	A	1
E	M	R	M	1
E	M	O	A	1
E	M	O	M	1
E	S	I	A	1
· · ·				
N	O	I	A	12

Pseudocode

The development of the CCT code involves four specific steps; each will be discussed below.

STEP 1 - DATA PREPARATION

- Pass each service or enrollment record for clients in the study population.
- Exclude records that do not represent service/enrollment during the study window.
- Check relevant records for appropriate service categories, service dates, enrollment categories, and enrollment dates.
- Fix missing or inappropriate dates
 - Impute missing dates or reassign date values if part of a date span is outside the study window
 - Reassign date values in cases where end dates come before begin dates.
- Sort records by client ID.

STEP 2 - COMPILE CLIENT SUMMARY RECORDS

- Pass records prepared in Step 1.
- When processing the first record for each client, initialize CTL variable to all dashes, and set client-level earliest and latest service dates to missing. Set earliest and latest enrollment dates to missing. Note that these variables will be retained across the processing of all records for each client.
- For each record, calculate the appropriate codes to be written to the segment(s) of the CTL variable that correspond to the date(s) on the service/enrollment record. Note that if more than one record relates to a given date segment on the CTL variable, the coding will represent the cumulative information from all processed records associated with that date.
- Update retained date variables (e.g., earliest and latest service dates), if appropriate.
- Upon completion of processing of the last record for a client, output one client summary record, retaining only client-level variables.

Note: If a study is designed to be episode-based, this step can be modified to define episodes and compile quantitative and qualitative information about each client episode. The output of summary records could be at the client-level or at the client-episode-level.

STEP 3 - COMPILE CLIENT-LEVEL SERVICE/ENROLLMENT RATES FOR NARROW TIME PERIODS AND SUCCESSIVELY AGGREGATE THESE INTO BROADER TIME PERIODS

- Pass client-level summary records from Step 2.
- Apply an algorithm that interrogates the CTL variable and compiles counts of unique dates, by selected service/enrollment profiles, over narrow time periods (e.g., week, month, quarter). An

example of such an algorithm might consist of, the number of outpatient encounter dates, in a given month, for MH treatment reported by Medicaid and during Medicaid enrollment.

- Create new variables that successively aggregate counts compiled for narrow time periods into broader time periods (e.g., months into quarters, quarters into years, and years into full study window).
- Develop qualitative variables (i.e., flag variables) from the quantitative counts of unique encounter dates.

Note: These flags should be [multi-positional] strings where each position represents a qualitative code to represent service/enrollment characteristics during a narrow time period, and the whole string represents a profile of service/enrollment characteristics over a broader unit of time. For example, a flag for quarterly MH inpatient treatment under Medicaid would have four positions; one for each quarter, where a particular code would represent any such service during a quarter and a dash would represent no such service during a quarter. In this example, 'IIII' would indicate at least one MH inpatient Medicaid service encounter during each quarter of a year; 'I-I-' would indicate such service encounters only in the first and third quarters; '----' would indicate no such service encounter during an entire year.

- Output one record per client that contains these new client-level variables, in addition to those created in Step 2.

STEP 4 (OPTIONAL) - EXAMINE FULL CTL VALUES FOR SELECTED CLIENTS

Apply an algorithm that prints out entire monthly segments of the CTL string, over the entire study window, for selected clients. This can be particularly useful for clients that appear to be outliers for certain measures, or who otherwise appear to be anomalous.

Note: For episode-based studies, these prints of the CTL can be very useful in fine-tuning the criteria for defining an episode.

Sample print of a CTL value for one client

Display of Monthly Service/Eligibility Patterns: Sequential date codes (5 bytes per day)

Byte 1: Agency MH SrvCode, Byte 2: Agency SA SrvCode, Byte 3: Medicaid EligCode, Byte 4: Medicaid MH SrvCode, Byte 5: Medicaid SA SrvCode

Key:

Medicaid EligCodes: D = Medicaid Disabled Status, E = Other Medicaid Eligibility, N = Not Medicaid Eligible

MH SrvCodes: I = Inpatient, R = Residential, O = Outpatient, W = R+O, X = I+O, Y = I+R, Z = I+R+O, - = None

SA SrvCodes: i = Inpatient, r = Residential, o = Outpatient, w = r+o, x = i+o, y = i+r, z = i+r+o, - = None

CLIENTID=ST00000002

[illegible][illegible][illegible]

Studies based on IDB data using Chronology-Compilation-Tool

- Bartosch W, Vandivort-Warren R, Dilonardo J, Schroeder D, Whalen D, Forhan C, Miller K, Federman E, Bray J (2007) Services for persons with evidence of withdrawal: an analysis of integrated state data. *in preparation*
- Bray JW, Vandivort-Warren R, Dilonardo J, Dunlap L, Schroeder D, Forhan C, Miller K (2007) Healthcare utilization of individuals with opiate use disorders: an analysis of integrated Medicaid and state mental health/substance abuse agency data. *Submitted to J Behav Health Serv Res.*
- Bray JW, Davis KL, Graver L, Schroeder D, Buck JA, Dilonardo J, Vandivort-Warren R (2005) Mental health and substance abuse treatment utilization among individuals served by multiple public agencies in 3 states. *J Behav Health Serv Res* 32:282-293.
- Buck JA, Teich JL, Graver L, Schroeder D, Zheng D (2004) Utilization of public mental health services by adults with serious mental illness. *Adm Policy Ment Health* 32:3-15.
- Buck JA (2001) IDB Project: Clients with Mental Disorders. Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Chalk M (2001a) Continuing Treatment Following Inpatient Care. Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Chalk M (2001b) IDB Project: Clients with Substance Use Disorders. Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Coffey RM, Graver L, Schroeder D, Busch J, Dilonardo J, Chalk M, Buck JA (2001) Mental Health and Substance Abuse Treatment: Results from a Study Integrating Data from State MH, SA, and Medicaid Agencies, SAMHSA Publication number SMA-01-3528. Rockville, MD: Center for Substance Abuse Treatment and Center for Mental Health Services, Substance Abuse and Mental Health Services Administration
<http://www.csat.samhsa.gov/IDBSE/idb/reports/FinalReportv9.pdf>

- Dilonardo J (2001a) IDB Project: Clients with Mental and Substance Abuse Disorders. Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Dilonardo J (2001b) Receipt of Mental Health and Substance Abuse Treatment Services by Persons with MH & SA Diagnoses. In: Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Federman EB, Bray J, Whalen D, Miller K, Forhan C, Buck J, Vandivort-Warren R (2007) Identifying opportunities to maximize Medicaid matching funds: An analysis of the integrated Medicaid and state MH/SA agency data. *in preparation*
- Larson M, Zhang A, Schroeder D, Tompkins C, Reif S, Schmidt L (2007) Measuring Medicaid substance abuse services performance quality: New knowledge combining Medicaid and state agency data. *in preparation*
- Lloyd J (2001) Expenditures and Utilization of Mental Health and Substance Abuse Services for Youth in Oklahoma and Washington. In: Presented at: Integration of State-Level Mental Health, Substance Abuse, and Medicaid Data. Sponsored by CSAT/CMHS. Bethesda, MD, July 27.
- Mark TL, Vandivort-Warren R, Montejano LB (2006) Factors affecting detoxification readmission: analysis of public sector data from three states. *J Subst Abuse Treat* 31:439-445.
- Teich JL, Buck JA, Graver L, Schroeder D, Zheng D (2003) Utilization of public mental health services by children with serious emotional disturbances. *Adm Policy Ment Health* 30:523-534.

References

- Buck JA, Teich JL, Graver L, Schroeder D, Zheng D (2004) Utilization of public mental health services by adults with serious mental illness. *Adm Policy Ment Health* 32:3-15.
- Federman EB, Bray JW, Whalen D, Miller K, Forhan C, Buck J, Vandivort-Warren R (2007) Identifying opportunities to maximize Medicaid matching funds: An analysis of the integrated Medicaid and state MH/SA agency data. *in preparation*
- Mark TL, Vandivort-Warren R, Montejano LB (2006) Factors affecting detoxification readmission: analysis of public sector data from three states. *J Subst Abuse Treat* 31:439-445.